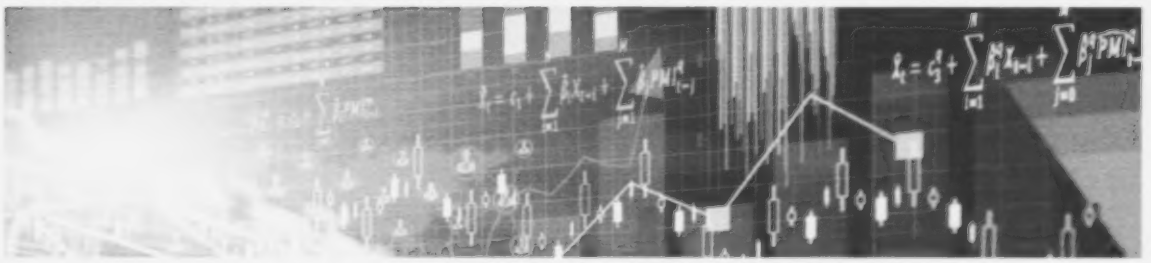
BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2017-47

Cash Versus Card: Payment Discontinuities and the Burden of Holding Coins



by Heng Chen, Kim P. Huynh and Oz Shy

Bank of Canada staff working papers provide a forum for staff to publish work-in-progress research independently from the Bank's Governing Council. This research may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2017-47

November 2017

Cash Versus Card: Payment Discontinuities and the Burden of Holding Coins

by

Heng Chen,[1] Kim P. Huynh[1] and Oz Shy[2]

[1] Currency Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
hchen@bankofcanada.ca
khuynh@bankofcanada.ca

[2] ozshy@ozshy.com

ISSN 1701-9397

© 2017 Bank of Canada

Acknowledgements

This paper was started while Oz Shy was visiting the Bank of Canada in May 2016. We thank Victor Aguirregabiria, Hiroshi Fujiki, Ben Fung, Ted Garanzotis, Charles Kahn, Bill Laur, Thomas Lemieux, Angela Redish and Richard Wall for helpful discussions and suggestions. We also benefited from participants at the Bundesbank conference War on Cash: Is there a Future for Cash? Island of Mainau, Germany, April 2017. We thank Rallye Shen for excellent research assistance on early versions of this paper. The views expressed in this article are those of the authors and do not necessarily represent the views of the Bank of Canada.

Abstract

Cash is the preferred method of payment for small value transactions generally less than $25. We provide insight to this finding with a new theoretical model that characterizes and compares consumers' costs of paying with cash to paying with cards for each transaction. Our novel method accounts for how much change is received in the form of banknotes and metal coins, assuming that the weight and size of coins are inconvenient to carry. We use the regression discontinuity design (RDD) approach to estimate the model using the 2013 Bank of Canada Method-of-Payments (MOP) Survey and find a significant number of cash users who switch to paying with debit or credit cards at transaction values marginally above $5 and $10. We attribute this finding to the burden of receiving coins as change associated with the currency denomination structure. Our proposed methodology is general and can be applied to other countries and institutional details.

Bank topics: Bank notes; Econometric and statistical methods

JEL codes: D03, E42

Résumé

L'argent comptant est le mode de paiement privilégié pour les petites transactions qui en général ne dépassent pas 25 $. Nous apportons des éclaircissements sur ce point à l'aide d'un nouveau modèle théorique qui permet de définir et de comparer, pour chaque transaction, les coûts associés au paiement en argent comptant par rapport à ceux qui sont rattachés au paiement par carte. Notre modèle inédit permet de déterminer à combien s'élève la petite monnaie reçue sous forme de billets et de pièces, en partant du principe que ces dernières sont peu pratiques vu leur poids et leur taille. Nous adoptons l'approche de discontinuité de la régression pour estimer le modèle à partir de l'enquête sur les modes de paiement réalisée en 2013 par la Banque du Canada. D'après nos conclusions, un grand nombre d'utilisateurs d'argent comptant paient par carte de débit ou de crédit dès que la valeur des transactions dépasse légèrement 5 $ et 10 $, respectivement. Cette observation s'explique par le fardeau associé au fait de recevoir de la petite monnaie en pièces ainsi que par la nature même des dénominations monétaires. Notre proposition de méthode a une portée générale et peut s'appliquer à d'autres pays et dispositifs institutionnels.

Sujets : Billets de banque; Méthodes économétriques et statistiques

Codes JEL : D03, E42

Non-Technical Summary

Motivation and Question

This paper explores the hypothesis that consumer choice between cash and payment cards (debit or credit) is driven by the burden of holding coins. We answer this question by examining and quantifying the different holding attributes of cash in the forms of banknotes and metal coins.

Methodology

By exploiting discontinuities around thresholds from banknote denominations in Canada, i.e., $5, $10, and $20, we can treat each transaction as a quasi-natural experiment and apply the regression discontinuity design (RDD) to understand consumer payment choice. The RDD methodology addresses the endogeneity of transaction value and the choice of cash versus payment cards.

Key Results and Contributions

Results confirm the hypothesis that consumers' cost of receiving coins as change affects their choice between paying with cash or with a card, and we empirically measure such cost from discontinuities in consumer payment choice. We find that the probability of using banknotes (relative to paying with debit or credit cards) is higher when the transaction generates a lower number of coins received as change, which occurs at $5 and $10. We also find that the probability of using banknotes falls by 19.2 and 11.9 percent at $5 and $10 transaction values.

1. Introduction

Cash use at the point-of-sale (POS) remains strong in most countries. A recent study by Bagnall et al. (2016) compares payment diaries from seven countries and shows that cash still dominates low-value transactions, especially those lower than $25. Using payment and scanner data, Bounie and Francois (2006), Klee (2008) and Fujiki and Tanaka (2017) document that transaction value is a strong determinant of payment choice. Further studies by Schuh and Stavins (2010), Simon, Smith, and West (2010), and Arango, Huynh, and Sabetti (2015) confirm this finding using payment diary data. Wang and Wolman (2016) illustrate this finding using extensive scanner data from the United States.

To illustrate this point, we use data from the 2013 Bank of Canada Method-of-Payments (MOP) Survey. Figure 1 displays a kernel density of transaction values and shows that most transactions are less than $25.[1] Further, Figure 2 shows the ratio of total cash transaction values to total cash and card transaction values as a function of transaction value. These figures illustrate that cash dominates for low value transactions. Studies by Kalckreuth, Schmidt, and Stix (2014) and Wakamori and Welte (2017) argue that preferences or habits are a major reason why consumers still use cash for lower-value transactions. Huynh, Schmidt-Dengler, and Stix (2014) discuss the role of the lack of universal acceptance of cards as a reason for the ongoing usage of cash. Fung, Huynh, and Stuber (2015) provide a high-level discussion of cash usage in Canada and document the fact that cash is used for small-value transactions and highlight the role of cost and security.[2]

All the papers cited above analyze the cash payment in the bundle of banknotes and coins; however, they do not differentiate various holding attributes between banknotes and coins. Here, banknotes are generally known as paper-money or paper notes, but also include cotton or polymer bills. This paper will look at whether payment choices between cash and cards are driven by the burden of holding coins.

This paper constructs a theoretical model that characterizes the choice of paying with cash and cards (both debit and credit card). The theoretical model predicts that consumers use cash at

[1] Detailed information about the 2013 MOP Survey is provided by Henry, Huynh, and Shen (2015).

[2] Chen, Felt, and Huynh (2017) document that contactless credit cards have a small effect on cash usage using longitudinal survey data from 2010 to 2012.

lower-value transactions because it is convenient, but this convenience is limited due to the weight of coins they receive as change. The major inconvenience of cash use, the need to carry metal coins in the wallet remained virtually unchanged over hundreds or even thousands of years.[3] A recent example is the introduction of the dollar coin in the United States in 2001. According to the report of the Government Accountability Office (2002) on page 21, the US Mint "*... further noted that consumer associations said the coin would be bulky and would add weight to wallets and pockets.*" Therefore, our approach adds a new dimension to the literature on monetary equilibria with discrete monetary units such as Wallace and Zhu (2004), Redish and Weber (2011) and Huang and Igarashi (2013). This new dimension is introduced by adding the burden of holding coins into consumers' payment utility functions.

To empirically estimate the burden of holding coins, we exploit thresholds of transaction values at \$5, \$10, \$15, \$20 and \$25. This is because our model provides a theoretical underpinning for discontinuity in the cash payment, where the probability of paying with cash conditional on transaction value is discontinuous at thresholds, due to the denomination structure of banknotes.[4] As an illustrative example, we assume that consumers receive change in the most efficient manner as to minimize the number of coins. That is, the change is composed of coins with the highest possible denomination so that a cash payment of \$4.95 with one \$10 banknote would result in one nickel (0.05) and a \$5 note. A cash purchase of \$5.05 would result in receiving six coins as change: two two-dollar (2.00), one fifty-cent (0.50), one quarter (0.25), and two dimes (0.10). Since consumers care about weights of their wallets and dislike change in the form of coins, the probability of using cash is higher to the left of \$5 and lower to the right, provided that other factors affecting payment choices are varying smoothly with respect to transaction values around \$5. Thus, cash payment discontinuities are driven by discontinuous numbers of coins received as change around thresholds, which are induced by the denomination structure. Table 1 displays coins and banknotes in circulation in Canada (Canadian dollars), Europe (euros), and the United States (dollars).

[3] Sargent and Velde (2002) provide a historical account of the difficulties of monetary authorities to provide sufficient amount of coinage for change. The authors coin this issue as "the big problem of small change," which leads many economies to undertake monetary reform and introduce fiat currency via banknotes.

[4] Threshold \$15 can be generated by one \$5 and one \$10 banknote, or three \$5 banknotes. Threshold \$25 can be generated similarly.

The previous literature on payment choices treats transactions as exogenous, which rules out the possibility of consumers choosing what to buy, when and where to purchase. However, by exploiting discontinuities around thresholds, we can treat each transaction as a quasi-natural experiment; see Imbens and Lemieux (2008). We apply the regression discontinuity design (RDD) method to quantify the casual effect of receiving coins as change on the payment choices by testing the validity of the RDD following Canay and Kamat (2016). In order to have a clear identification, this paper focuses on the sharp RDD by restricting transaction records where there are no coins to start with, in which coins in the cash holding are only from receiving coins as change, rather than coins accumulated as change from previous transactions.[5] We find that the probability of using banknotes (relative to paying with debit or credit cards) is higher when the transaction generates a lower number of coins received as change, which occurs at \$5 and \$10. We find that the probability of using banknotes falls by 19.2 and 11.9 percent at \$5 and \$10 transaction values.

The "burden of holding coins" on consumers can be investigated using two methods: First, as a static problem where the burden of receiving and carrying coins is repeated with each purchase. Using the static approach, each purchase can be modeled as a separate static problem. Second, as a dynamic optimization problem where the consumers' goal is to minimize the accumulation of coins by the end of a certain period; for example, by the end of the day. Although the dynamic optimization approach is an interesting problem, we choose to construct a static model because it captures the repeated burden of receiving coins as change faced by cash uses. As we show below, the static model is ideal for identifying the discontinuities with respect to transaction values when shifting between paying with cash and paying with a card. To be able to compare the transaction values where the discontinuities are realized, we need to model consumers who treat each transaction separately regardless of whether a transaction happens to be the first, second, or the third transaction in a given time period.

Our study is organized as follows. Section 2 constructs a consumer payment choice model based on consumers' costs and benefits of paying cash or with cards. Section 3 provides descrip-

[5]This paper studies the burden of holding coins by looking at people trying to avoid receiving coins from their transaction ex-post. And an alternative approach would be ex-ante, where they try to dump the accumulated coins from the previous transactions for the current purchase. The latter is left for the future research.

tive statistics from the 2013 Bank of Canada Method-of-Payments Survey. Section 4 develops and discusses the RDD estimation methods and discusses the results. Section 5 concludes.

2. A model of payment choice

2.1 Consumers' decision making process

The introductory section described several factors (such as cost, security, and other characteristics) that influence consumer payment choice. Let $v > 0$ denote the transaction value in terms of the tax-inclusive price of the good or service that the consumer would like to pay for. Here v is rounded to the nearest $5¢$ such as $v \in \{0.05, 1.95, 22.15\}$ and so on, because the $1¢$ coin in Canada has been eliminated and the price is rounded to the nearest $5¢$.

The model is based on the following reservation-price strategy S that specifies a threshold transaction value $\bar{v} > 0$ for which

$$S(v) \overset{def}{\begin{cases} H\ (cash) & \text{if } v \in [\bar{v} - \epsilon,\ \bar{v}] \\ D\ (card) & \text{if } v \in (\bar{v},\ \bar{v} + \epsilon], \end{cases}} \tag{1}$$

where ϵ is a small number. That is, under this strategy, transaction values on the left of $\bar{v}$ would be more likely to be paid with cash than on the right, where $\bar{v}$ are thresholds implied by the denomination structure of the banknotes.

The goal of the paper is (a) to examine whether the discontinuities of paying with cash conform to these thresholds, and (b) to conduct an econometric analysis to estimate the likelihood differences of paying with cash versus cards around the threshold values for average Canadian consumers.[6]

2.2 The cost of receiving change in metal coins

We model consumers who dislike carrying a "thick" or "heavy" wallet. Therefore, the strategy (1) takes into consideration the amount of change a consumer will receive in the form of metal coins for each transaction. The strategy then specifies whether the consumer will decide to pay with a

[6] In the RDD identification section, the representative agent model can be relaxed to accommodate for unrestricted heterogeneity of the payment choice, as long as there are common thresholds to generate the burden of holding coins. Such generalization is possible because the denomination scheme is binding.

card or with cash based on this amount, and based on other costs of using cash and cards.

We assume that banknotes are weightless. The left column in Table 1 displays the types of coin in circulation in Canada. We divide this list into two groups: Coins denominated by $5¢$, $10¢$, $25¢$, and $50¢$. These coins have value of less than $1 and will be referred to as *small coins*. The $1 and $2 coins will be referred to as *large coins*.

To measure the disutility (cost) from carrying a heavy wallet, let δ_S and δ_L denote a consumer's disutility from receiving change in the form of a small or a large coin, respectively. Note that the relationship between δ_S and δ_L may vary from one person to another. Consumers who are irritated by small coins will experience $\delta_S > \delta_L$. Consumers who are irritated by the total weight of a coin are captured by $\delta_S < \delta_L$.

Figure 3 displays the number of small coins received as change for all transaction values v between $0 and $1 (recall that $5¢ is the smallest coin in Canada, so all cash transactions are rounded accordingly). The underlying assumption behind Figure 3 is that consumers carry only large coins ($1 and $2) and banknotes. Let $i \in \mathbf{Z}_+ = \{0, 1, 2, \ldots\}$ be any nonnegative integer representing the dollar component of each transaction value, such as $0, $1, and $2 . A formal representation of Figure 3 is given in the following cost function:

$$c_S(v) = \begin{cases} 0 & \text{if } v = i \times \$1.00 \\ \delta_S & \text{if } v \in \{\$0.50, \$0.75, \$0.90, \$0.95\} + i \times \$1.00 \\ 2\delta_S & \text{if } v \in \{\$0.25, \$0.40, \$0.45, \$0.65, \$0.70, \$0.80, \$0.85\} + i \times \$1.00 \\ 3\delta_S & \text{if } v \in \{\$0.15, \$0.20, \$0.30, \$0.35, \$0.55, \$0.60\} + i \times \$1.00 \\ 4\delta_S & \text{if } v \in \{\$0.05, \$0.10\} + i \times \$1.00. \end{cases} \tag{2}$$

We now provide few examples that illustrate how the small coins cost function (2) and Figure 3 are constructed. First, all transactions with even dollars ($v = i \times \$1$, $i \in \mathbf{Z}_+$) are paid for in full with large coins or banknotes, so no small change is returned to the consumer.

Suppose the consumer pays $v = \$0.90$ using a $1 large coin. Then, the consumer receives $0.10 change back (one small coin). For a payment of $v = \$0.80$ the consumer receives two small coins as change. For a payment of $v = \$0.35$, the change that minimizes the number of small coins consists of three coins: $0.05, $0.10, and $0.50. Finally, for a payment of $v = \$0.05$, the change that minimizes the number of small coins consists of four coins: two of $0.10, plus $0.25 and $0.50.

Moving on to change consisting of large coins, suppose that consumers carry only $5 banknotes and up, and no coins. Figure 4 illustrates a cycle of $5 where the number of large coins declines to zero for transaction value close to $v = \$5$, where a payment of $v = \$5$ can be fully paid with a banknote, so no change in coins is needed. A formal representation Figure 4 is given in the following cost function:

$$c_L(v) = \begin{cases} 0 & \text{if } \$4.05 \leq v \leq \$5.00 + i \times \$5.00 \\ \delta_L & \text{if } \$2.05 \leq v \leq \$4.00 + i \times \$5.00 \\ 2\delta_L & \text{if } \$0.05 \leq v \leq \$2.00 + i \times \$5.00. \end{cases} \tag{3}$$

To understand how the large coins' cost function (3) and Figure 4 were constructed, we examine some representative transaction values. First, all transactions with multiples of five dollars ($v = i \times \$5$, $i \in \mathbf{Z}_+$) are paid for in full with banknotes, so no large coins are needed.

If the consumer pays with a $5 bill to pay for a transaction value of $0.05, the consumer will receive two $2 large coins in addition to small coins. Next, paying $2.05 with a $5 bill will return one $2 large coin in addition to small coins. Similarly, paying $3.05 with a $5 bill will return one $1 large coin in addition to small coins. Finally, any transaction exceeding $4.05 with a $5 bill will not return any large coins (only small coins).

2.3 Cost of card payments and total cost of making a payment

Suppose that consumers bear cost of paying with cards $c_D(v)$, which declines with the value of transactions so that $\Delta c_D(v)/\Delta v < 0$. This cost element reflects a preference for banknotes over cards for low-value transactions. Large-value transactions have low cost of paying with cards, reflecting also increased benefits from card rewards, which are proportional to the transaction value.

To sum up, given a transaction value v, a consumer's decision of whether to pay cash or with a card is guided by comparing the cost of cash with the cost of cards to solve

$$\min_{H,D} C(v) = \begin{cases} H & \text{if } c_S(v) + c_L(v) \leq c_D(v) \\ D & \text{if } c_S(v) + c_L(v) > c_D(v). \end{cases} \tag{4}$$

From this payment cost minimization problem, we want to extract the threshold transaction values

r that define consumers' payment choice strategy defined in (1).

Figure 5 illustrates total cost of paying cash defined in (2) and (3), $c_S(r) + c_L(r)$ (see first row in (4)), for the special case where $\delta_S = \delta_L = 1$. In this case, consumers assign equal disutility to a large and small coin (normalized to one). The domain of Figure 5 is restricted to transaction values up to \$25. Note that this cycle of the total cost of paying with cash repeats itself every \$5.

3. 2013 Method-Of-Payments Survey

The 2013 Method-Of-Payments (MOP) survey is a payment diary survey commissioned by the Bank of Canada. Further details of the results of this survey are described in the report by Henry, Huynh, and Shen (2015). Chen et al. (2016) provides a discussion of the statistical methodology employed to ensure that survey is statistically representative.

In order to apply the RDD, we define the outcome variable as an indicator of whether the consumer uses cash for the purchase or not. The binary treatment is the burden of holding coins, where the transaction would result in many coins at the transaction values marginally above \$5, \$10, \$15, \$20 and \$25. In order to be able to test the theoretical model discussed in Section 2, we restrict the sample to:

1. Transactions that allow for a choice between cash or card payment.

2. Individuals with accurate information on cash identity error within (+/- 5 CAD) [7], so that the precision of data is ensured without serious measurement error.

3. Transactions where the respondent made neither cash withdrawal nor other cash managements (e.g., deposit cash and give cash as gifts). Such a restriction is to obtain an accurate amount of cash held before the purchase. In the 2013 MOP, respondents recorded purchases and cash withdrawals in separate sections of the diary, so it is only possible to establish the sequence of purchases with respect to one another, and not in relation to the withdrawals.

[7] In order to gauge data quality, Henry, Huynh, and Shen (2015) use the cash identity error, which is the deviation between how much cash they have left at the end of three days' diary and the sum of the amount of cash they have at the start of the diary, and how much cash they spend/receive during the three days of completing the diary. The amount of error in this identity therefore provides a measure of how accurately the diary was completed, similar to how a household might balance its chequebook. They find that 80 percent of diaries have almost perfect cash identity.

4. Only the first transaction of the three-day diary, because the 2013 MOP diary only asks about the number of banknotes and coins at the beginning of the three-day diary. By focusing on the first transaction of the three-day diary, we can obtain the precise configuration of banknotes and coins in people's wallets. In addition, we further restrict the sample of the first transaction to the sub-sample of no coins held in their wallets.

The above restrictions leave us with 518 transaction records that can be used in our RDD estimation.

4. Empirical testing using RDD

Following Angrist and Lavy (1999), we use the RDD to confirm discontinuities ("big jumps") in cash usage around transaction values of $5, 10, 15, 20$ and 25 dollars. Although the theory in Section 2 describes many "small jumps" at each five-cent increment, we will focus only on the five steepest jumps marginally above the transaction values at $5, 10, 15, 20$ and 25 dollars. The main reason that we focus only on these five steepest jumps is because identifying/estimating small jumps is very difficult, especially when jump locations are too close (Mallat (1999)). As an approximation, we are ironing out these small discontinuities within each five-dollar interval, so that the relationship between the total number of coins received and transaction values are piecewise linear downward sloping lines, with jumps at knots at $5, 10, 15, 20$ and 25 dollars; see red solid lines in Figure 5. Hence, we define $B = I\{v > \bar{v}\}$ to be the binary treatment variable that takes a value of 1 for transactions that are affected by the burden of holding coins and 0 for transactions that are not, where $\bar{v} \in \{5, 10, 15, 20, 25\}$.

Consider $Y = S(B, v, u)$ where Y is the binary dependent variable where 1 is paid with cash and 0 with card, and u is an unobserved covariate vector. Since we analyze transactions in which consumers have only banknotes, $Y = 1$ indicates transactions paid with banknotes. Notice that v is generally correlated with u, because consumers usually choose what to buy, when and where to make transactions, which makes v, and therefore B, endogenous. This nonseparable model with endogenous variables v and B makes it difficult to identify S and the conditional distribution of u on v; see Chesher (2003) and Matzkin (2007). However, the treatment B is discontinuous at

thresholds $\widetilde{v}$, thus by Lee (2008) we can identify

$$\lim_{v \downarrow \widetilde{v}} E(Y|v) - \lim_{v \uparrow \widetilde{v}} E(Y|v) \tag{5}$$

as the weighted average treatment effect (ATE) at $\widetilde{v}$

$$\frac{1}{f_v(\widetilde{v})} E_u \left[f_{v|u}(\widetilde{v}|u) \left(g_1(\widetilde{v}, u) - g_0(\widetilde{v}, u) \right) \right] \tag{6}$$

where f_v is the density of v, $Y_1 = g_1(v, u)$ and $Y_0 = g_0(v, u)$ are potential banknote-payment choices affected by the burden of holding coins or not.

Heuristically, the discontinuity of the conditional mean Y on v can be derived based on total cost (4), where $B = I\{c_S(v) + c_L(v) - c_D(v) > 0\}$ and $S = I\{B \leq u\}$. We observe that the conditional expectation of Y on v is

$$E(Y|v) = \Pr\left[I\{c_S(v) + c_L(v) - c_D(v) > 0\} \leq u|v \right]. \tag{7}$$

Given that $c_D(v)$ is continuous for all v[5], while $\lim_{v \downarrow \widetilde{v}} |c_S(v) + c_L(v)| - \lim_{v \uparrow \widetilde{v}} |c_S(v) + c_L(v)| \neq 0$ at $v = \widetilde{v} \in \{5, 10, 15, 20, 25\}$, we have

$$\lim_{v \downarrow \widetilde{v}} E(Y|v) - \lim_{v \uparrow \widetilde{v}} E(Y|v) \neq 0 \tag{8}$$

at $v = \widetilde{v} \in \{5, 10, 15, 20, 25\}$. The graphical RDD estimates in Figure 6 exhibit negative jumps around thresholds in terms of probabilities of banknotes payments.

In this paper, we choose the sharp RDD approach to quantify the effects of receiving coins as change on the payment choice, where the burden of holding coins dummy variable B is a deterministic function of transaction values v. Such a deterministic relationship between B and v assumes that every time the transaction is carried out, the number of coins received satisfies (2) and (3), and thus Figure 5. If consumers receive change that is different from what the model predicts, then the fuzzy RDD approach seems more appropriate. However, we choose to adhere to the sharp RDD for two reasons. First, our sharp RDD provides a lower bound estimate, given that the denominator of the fuzzy RDD estimate is the probability of facing burdens of holding

[5]Many other factors (risk and security preferences, etc.) could affect the payment choice, but they are all absorbed into c_D, as long as their influences vary smoothly with respect to v.

coins, which is always smaller or equal to 1. Second, the 2013 Bank of Canada MOP does not ask how many coins were received after each transaction, thus the number of coins has to be imputed by (2) and (3).

4.1 Sharp RDD estimates

Following Lee (2008), we obtain the linear auxiliary regression:

$$Y = g(v) + \delta B + \varepsilon, \tag{9}$$

where the unknown nonparametric function $g(\cdot)$ is continuous on the support of v, $E(\varepsilon|v) = 0$ and

$$\delta = \lim_{v \downarrow \bar{v}} E(Y|v) - \lim_{v \uparrow \bar{v}} E(Y|v). \tag{10}$$

However, recall that the transaction values v are rounded to the nearest $5¢$ because of the elimination of the $1¢$ coin in Canada, thus v is a discrete variable by nature. This makes it impossible to compare outcomes for "just above" and "just below" the threshold, which requires us to choose a specific functional form for the relationship between v and Y. We will restrict $g(v)$ to the parametric polynomial specification, which leads to

$$Y = \beta_0 + \beta_1 v + \beta_2 v^2 + \delta B + \varepsilon, \tag{11}$$

where $g(v) \equiv \beta_0 + \beta_1 v + \beta_2 v^2$. However, the parametric functional assumption will create specification error between Y and v, thus we follow Lee and Card (2008) to use the clustered standard errors. In addition, to assess the credibility of the RDD, we test the continuity of the distribution of baseline covariates at the thresholds using Canay and Kamat (2016), which is a suitable procedure for discrete running variable v. Table 2 shows that our RDD setup is indeed valid.[9]

In this paper, we estimate the causal effects of the burden of holding coins on banknote-

[9] However, due to the discreteness of v, the density test proposed by McCrary (2008) is infeasible.

payments based on

$$
\begin{aligned}
Y &= \beta_0 + \beta_1 v + \beta_2 v^2 + \delta_5 I\{v > 5\} + \alpha X + \varepsilon, \\
Y &= \beta_0 + \beta_1 v + \beta_2 v^2 + \delta_{10} I\{v > 10\} + \alpha X + \varepsilon, \\
Y &= \beta_0 + \beta_1 v + \beta_2 v^2 + \delta_{15} I\{v > 15\} + \alpha X + \varepsilon, \\
Y &= \beta_0 + \beta_1 v + \beta_2 v^2 + \delta_{20} I\{v > 20\} + \alpha X + \varepsilon, \\
Y &= \beta_0 + \beta_1 v + \beta_2 v^2 + \delta_{25} I\{v > 25\} + \alpha X + \varepsilon.
\end{aligned} \tag{12}
$$

An alternative specification is to include five thresholds into a single equation. This specification is estimated in order to capture spill-over effects across these thresholds:

$$
\begin{aligned}
Y &= \beta_0 + \beta_1 v + \beta_2 v^2 + \delta_5 I\{v > 5\} + \delta_{10} I\{v > 10\} + \delta_{15} I\{v > 15\} \\
&+ \delta_{20} I\{v > 20\} + \delta_{25} I\{v > 25\} + \alpha X + \varepsilon,
\end{aligned} \tag{13}
$$

where X is a vector of other observed covariates, such as, age, gender and banknotes holding before the transaction.[10] Notice that the left continuous function $I\{v > \tilde{v}\}$ is used instead of the right continuous function $I\{v \geq \tilde{v}\}$, because the burden of coins is minimized at $\tilde{v} \in \{5, 10, 15, 20, 25\}$; however, the burden of holding coins increases dramatically if v is only marginally above $\tilde{v}$.

4.2 Results

First, we apply the sharp RDD regressions (12) using ordinary least squares (OLS) for each threshold, separately. Table 3 illustrates the estimated marginal effects from the burden of holding coins (negative probabilities) at \$5, \$10, \$20 and \$25. The effects are largest and statistically significant for \$5 and \$10 with a marginal effect of -0.192 and -0.119. There is slight positive effect for \$15, but not statistically significant. Moreover, we also observe that these magnitudes of marginal effect estimates are decreasing between \$5, \$10, \$20 and \$25, which implies that burdens of holding coins are the largest at \$5 and \$10. Although the model predicts the same cost of holding coins, the diminishing (or insignificant) coefficients above \$10 might be because consumers' cost of cards

[10] Other observed covariates X are included to reduce sampling variability in the impact RDD estimates and is analogous to including baseline characteristics in the analysis of a randomized experiment.

$c_D(v)$ becomes more diverse as v increases: some have high-reward credit cards, while others do not. Thus, the expected burden of coins at \$15, \$20 and \$25 might be mitigated (or countered) by specific feature of cards. These results imply that at these \$5 and \$10 cutoff points, consumers get the most number/weights of coins, so that these dummies capture disutility of coins from the transaction. Therefore, the probability of using banknotes is reduced if the transaction value exceeds these cutoffs.

Second, we run the single OLS regression (13) that includes all the thresholds' dummies. The same regressors are used as in the above individual regressions. The single regression formulation can capture the joint dependence across thresholds, such as whether there is any substitution or complementarity among different denomination thresholds. Column (6) in Table 3 reveals that the estimated marginal effects are negative at \$5, \$10, \$20 and \$25 with the estimate at \$5 statistically significant. Notice that the estimated marginal effects in the single regression are larger than in the separate threshold regressions. Therefore, the findings from this single regression strengthen our results from individual threshold regressions for \$5, \$10, \$20 and \$25, and also indicates that there might be some substitution effect for the threshold at \$15.

Third, because the OLS estimates may range outside the range 0 and 1, we also use a single logit regression for all the thresholds combined, Column (7) of Table 3. In order to compare marginal effects, we compute the average partial effects for the logit regression, and find the sizes to be similar to the OLS linear model.

Furthermore, we run 25 individual threshold regressions where each threshold is determined at every one-dollar transaction value, where red dots in Figure 7 are coefficients of discontinuities in banknotes-use probabilities. Again, the peaks at $5-\epsilon$ and $10-\epsilon$ are still statistically significant, where peaks $15-\epsilon$, $20-\epsilon$ and $25-\epsilon$ are not. In addition, the figure displays drops at $5+\epsilon$ and $10+\epsilon$, and these troughs are consistent with our theory predication.

Figure 7 illustrates whether discontinuities in banknotes-use probabilities occur at other transaction values that may not correspond to the denomination structure banknotes. In the range between \$5 and \$10, coefficients of threshold at transaction values away from \$5 and \$10 are behaving as negative trend functions, which capture the downward-sloping relationship between

transaction values and banknotes-use probabilities as in Arango, Huynh, and Sabetti (2015). In addition, since there exist no other discontinuities in banknotes-use probabilities except $5 and $10, our smooth approximation within each five-dollar interval is justified.

5. Conclusion and implications for consumer payment choice

In this paper, we model consumers' cost of receiving coins as change, and empirically measure such cost from discontinuities in consumer payment choice. Our novel result adds a new perspective on consumers' payment choices between cash and card at low transaction values. Our analysis exploits the denomination thresholds to focus on payment choice *locally* at these low-value transactions. Compared to treating the low transaction values *globally*, which ignores the important role of coins, our RDD analysis identifies the *burden of coins* from cash payments at low transaction values. The methodology developed in the paper is sufficiently general and therefore can be applied to other countries and institutional details to understand the implication of potential changes in the denomination structure. Bagnall et al. (2016) provide a seven-country, cross-country comparison of consumer payment diaries and find that cash is the preferred method of payment at low-value transactions. The theoretical model can be amended to reflect the relevant denomination structure discussed in Table 1. In the Eurozone there is a wider variety of coins in circulation compared to Canada. More precisely, the Euro adds 1¢, 2¢, 10¢, and 20¢ (but does not have a 25¢) compared with Canada, a net addition of three small coins. These differences would require some modifications of Figure 3, Figure 4, and Figure 5 that compute a customer burden of receiving coins as change at the point of sale. The RDD methodology could then be applied to the consumer payment diary data for each respective country.

Finally, we acknowledge the differences between the static and dynamic approaches to modeling consumer payment choices at the point of sale. We argued that the static approach is more suitable for identifying the transaction values at which consumers' payment choice exhibits a discontinuity. A dynamic variant of the static model would be based on the assumption that consumers are annoyed by the accumulation of coins rather than just by having to handle and carry coins when received as change in each transaction. The dynamic approach models consumers

who may have some coins in their pockets when entering a new point of sale. In this case, consumers may be able to pay with coins accumulated from previous transactions. However, empirical testing would require additional data that is hard to obtain. More precisely, testing a dynamic model would require data on the specific number of coins and notes consumers have before and after each transaction they make during a given time period, such as a day. These data may be burdensome and expensive to obtain.

References

Angrist, Joshua and Victor Lavy. 1999. "Using Maimonides' Rule to estimate the effect of class size on scholastic achievement." *Quarterly Journal of Economics* 114 (2):533–575.

Arango, Carlos, Kim P. Huynh, and Leonard Sabetti. 2015. "Consumer payment choice: Merchant card acceptance versus pricing incentives." *Journal of Banking and Finance* 55 (0):130 – 141. URL http://www.sciencedirect.com/science/article/pii/S0378426615000321.

Bagnall, John, David Bounie, Kim Huynh, Anneke Kosse, Tobias Schmidt, Scott Schuh, and Helmut Stix. 2016. "Consumer Cash Usage: A Cross-Country Comparison with Payment Diary Survey Data." *International Journal of Central Banking* 12 (4):1–61. URL https://ideas.repec.org/a/ijc/ijcjou/y2016q4a1.html.

Bounie, David and Abel Francois. 2006. "Cash, Check or Bank Card? The Effects of Transaction Characteristics on the Use of Payment Instruments." Mimeo.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2015. "Optimal data-driven regression discontinuity plots." *Journal of the American Statistical Association* 110 (512):1753–1769.

Canay, Ivan A and Vishal Kamat. 2016. "Approximate permutation tests and induced order statistics in the regression discontinuity design." Tech. rep., cemmap working paper, Centre for Microdata Methods and Practice.

Chen, Heng, Marie-Hélène Felt, and Kim P. Huynh. 2017. "Retail payment innovations and cash usage: accounting for attrition by using refreshment samples." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 180 (2):503–530. URL http://dx.doi.org/10.1111/rssa.12208.

Chen, Heng, Chris Henry, Kim Huynh, Q Shen, and Kyle Vincent. 2016. "Measuring Consumer Cash Holdings: Lessons from the 2013 Bank of Canada Methods-of-Payment Survey." *Survey Practice* 9 (4). URL http://www.surveypractice.org/index.php/SurveyPractice/article/view/325.

Chesher, Andrew. 2003. "Identification in nonseparable models." *Econometrica* 71 (5):1405–1441.

Fujiki, Hiroshi and Migiwa Tanaka. 2017. "Choice of payment instrument for low-value transactions in Japan." URL https://ideas.repec.org/p/zbw/iccp17/162909.html.

Fulford, Scott L., Claire Greene, and William Murdock. 2015. "U.S. consumer holdings and use of $1 Bills." Research Data Report 15-1, Federal Reserve Bank of Boston. URL https://ideas.repec.org/p/fip/fedbdr/15-1.html.

Fung, Ben, Kim P. Huynh, and Gerald Stuber. 2015. "The Use of Cash in Canada." *Bank of Canada Review* 2015 (Spring):45–56. URL https://ideas.repec.org/a/bca/bcarev/v2015y2015ispring15p45-56.html.

Government Accountability Office. 2002. "New Dollar Coin: Marketing Campaign Raised Public Awareness but Not Widespread Use." Tech. Rep. GAO-02-896, Government Accountability Office. URL http://www.gao.gov/products/GAO-02-896.

Henry, Christopher, Kim Huynh, and Rallye Shen. 2015. "2013 Methods-of-Payment Survey Results." Discussion Papers 15-4, Bank of Canada. URL https://ideas.repec.org/p/bca/bocadp/15-4.html.

Huang, Pidong and Yoske Igarashi. 2013. "Why ten $1's are not treated as a $10." University of Exeter, Economics Department Discussion Paper No. 13/10.

Huynh, Kim, Philipp Schmidt-Dengler, and Helmut Stix. 2014. "The Role of Card Acceptance in the Transaction Demand for Money." Staff Working Papers 14-44, Bank of Canada. URL https://ideas.repec.org/p/bca/bocawp/14-44.html.

Imbens, Guido and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142 (2):615–635.

Kalckreuth, Ulf Von, Tobias Schmidt, and Helmut Stix. 2014. "Using Cash to Monitor Liquidity: Implications for Payments, Currency Demand, and Withdrawal Behavior." *Journal of Money, Credit and Banking* 46 (8):1753–1786. URL https://ideas.repec.org/a/wly/jmoncb/v46y2014i8p1753-1786.html.

Klee, Elizabeth. 2008. "How people pay: Evidence from grocery store data." *Journal of Monetary Economics* 55 (3):526–541. URL http://ideas.repec.org/a/eee/moneco/v55y2008i3p526-541.html.

Lee, David S. 2008. "Randomized experiments from non-random selection in US House elections." *Journal of Econometrics* 142 (2):675–697.

Lee, David S and David Card. 2008. "Regression discontinuity inference with specification error." *Journal of Econometrics* 142 (2):655–674.

Mallat, Stéphane. 1999. *A wavelet tour of signal processing*. New York: Academic Press.

Matzkin, Rosa L. 2007. "Nonparametric identification." *Handbook of Econometrics* 6:5307–5368.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of econometrics* 142 (2):698–714.

Redish, Angela and Warren E. Weber. 2011. "A model of commodity money with minting and melting." Staff Report 460, Federal Reserve Bank of Minneapolis. URL https://ideas.repec.org/p/fip/fedmsr/460.html.

Sargent, Thomas J. and Franois R. Velde. 2002. *The big problem of small change*. Princeton economic history of the Western world. Princeton, NJ [u.a.]: Princeton Univ. Press.

Schuh, Scott and Joanna Stavins. 2010. "Why Are (Some) Consumers (Finally) Writing Fewer Checks? The Role of Payment Characteristics." *Journal of Banking & Finance* 34 (8):1745–1758.

Simon, John, Kylie Smith, and Tim West. 2010. "Price Incentives and Consumer Payment Behaviour." *Journal of Banking and Finance* 34 (8):1759–1772. New Contributions to Retail Payments: Conference at Norges Bank (Central Bank of Norway) 14-15 November 2008.

Wakamori, Naoki and Angelika Welte. 2017. "Why Do Shoppers Use Cash? Evidence from Shopping Diary Data." *Journal of Money, Credit and Banking* 49 (1):115–169. URL http://dx.doi.org/10.1111/jmcb.12379.

Wallace, Neil and Tao Zhu. 2004. "A commodity-money refinement in matching models." *Journal of Economic Theory* 117 (2):246–258.

Wang, Zhu and Alexander L. Wolman. 2016. "Payment choice and currency use: Insights from two billion retail transactions." *Journal of Monetary Economics* 84 (C):94–115. URL https://ideas.repec.org/a/eee/moneco/v84y2016icp94-115.html.

Table 1: Metal coins and banknotes in circulation in Canada, Europe, and the U.S.

Canada (CAD)		Europe (EUR)		U.S. (USD)	
Coins	Notes	Coins	Notes	Coins	Notes
0.05	5	0.01	5	0.01	1
0.10	10	0.02	10	0.05	2
0.25	20	0.05	20	0.10	5
0.50	50	0.10	50	0.25	10
1.00	100	0.20	100	0.50	20
2.00		0.50	200	1.00	50
		1.00	500		100
		2.00			

Notes: Note that Canada and Europe do not use paper notes denominated as one or two dollars or euros. In contrast, in the U.S., repeated introductions of the $1 coin have always failed because users find the $1 bill more attractive; see Fulford, Greene, and Murdock (2015). Another difference is that the U.S. does not have a $2 coin probably because such a coin would fail to be adopted as users would prefer using two $1 paper bills, as $2 bills are rarely used.
In 2012, Canada eliminated the $0.01 or the penny:
http://www.mint.ca/store/mint/about-the-mint/phasing-out-the-penny-6900002.

Table 2: P-values from the RD distribution permutations test

Running variable	Covariates		
Transaction value	Age	Income	Cash holding before transaction
Threshold at 5.05	0.72	0.24	0.15
Threshold at 10.05	0.30	0.87	0.86
Threshold at 15.05	0.26	0.76	0.48
Threshold at 20.05	0.43	0.47	0.33
Threshold at 25.05	0.33	0.90	0.63

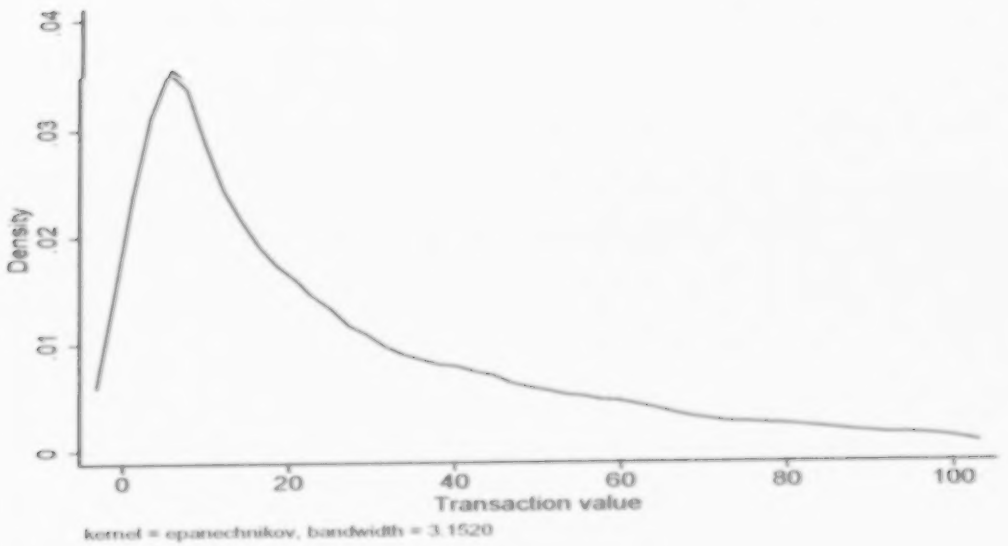
Notes: P-values from RD distribution tests using permutations are implemented by Stata command *rdperm* by Canay and Kamat (2016) to asses the credibility of the RDD. The null hypothesis is that the conditional distribution of the covariate on the transaction value is continuous at the threshold.

Table 3: Regression discontinuity design estimates of cash denomination thresholds

	Ordinary Least Squares						*Logit*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	0.627***	0.579***	0.580***	0.571***	0.577***	0.617***	0.499*
	(0.054)	(0.052)	(0.053)	(0.053)	(0.054)	(0.054)	(0.259)
Age between 35 and 53	−0.040	−0.043	−0.045	−0.045	−0.045	−0.089	−0.01
	(0.049)	(0.050)	(0.050)	(0.050)	(0.050)	(0.049)	(0.086)
Age over 55	0.162***	0.156***	0.160***	0.161***	0.160***	0.155***	0.19**
	(0.053)	(0.053)	(0.053)	(0.053)	(0.054)	(0.053)	(0.07)
Purchase amount	−0.011***	−0.012***	−0.016***	−0.013***	−0.015***	−0.007**	−0.01
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.01)
Purchase amount squared	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.00005***	0.00006*
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00002)	(0.00003)
Cash holding before transaction	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Threshold at $5	−0.192***					−0.194***	−0.205**
	(0.054)					(0.061)	(0.070)
Threshold at $10		−0.119*				−0.123	−0.128
		(0.065)				(0.087)	(0.102)
Threshold at $15			0.024			0.184*	0.204*
			(0.077)			(0.102)	(0.114)
Threshold at $20				−0.064		−0.148	−0.164
				(0.089)		(0.121)	(0.127)
Threshold at $25					−0.089	−0.062	−0.089
					(0.096)	(0.121)	(0.150)

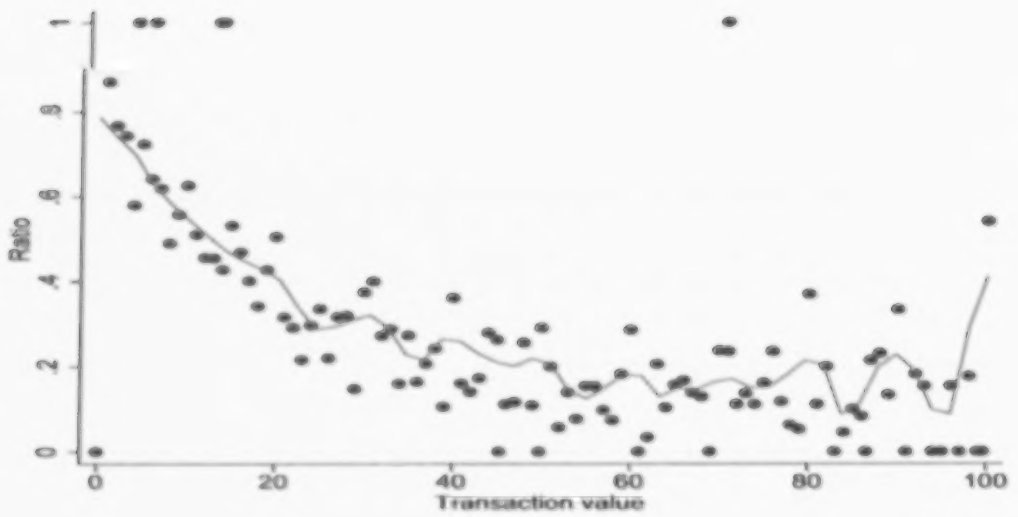
Notes: The dependent variable is the binary choice of whether to use cash or not. The clustered standard errors in brackets are reported using Lee and Card (2008) and *, **, and *** denote the marginal significance level of 10, 5, and 1 percent, respectively. There are 518 observations used in these regressions. For the logistic regression, the average partial effects are reported instead of coefficients, and standard errors of average partial effects are reported.

Figure 1: Distribution of transaction values



Notes: Kernel density is drawn for the transaction value less than $100.

Figure 2: Ratio of total cash transaction values to total cash and card values



Notes: The scatter plot (blue dots) is the ratio of total cash transaction values to total cash and card transaction values. The red curve is the nonparametric local linear fit for the percentage against the transaction value.

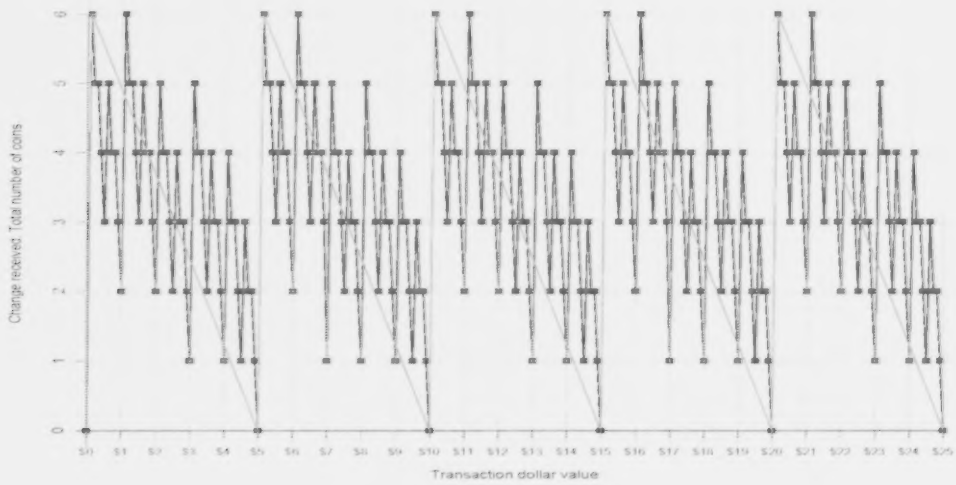
Figure 3: Number of small coins as change for transactions between zero and one dollar

Notes: The figure assumes that consumers carry only large coins ($1 and $2) and no small coins. Connecting lines do not reflect possible transactions.

Figure 4: Number of large coins received as change for transactions between $0.00 to $5.00, assuming that consumers carry only banknotes ($5 and up) and no coins

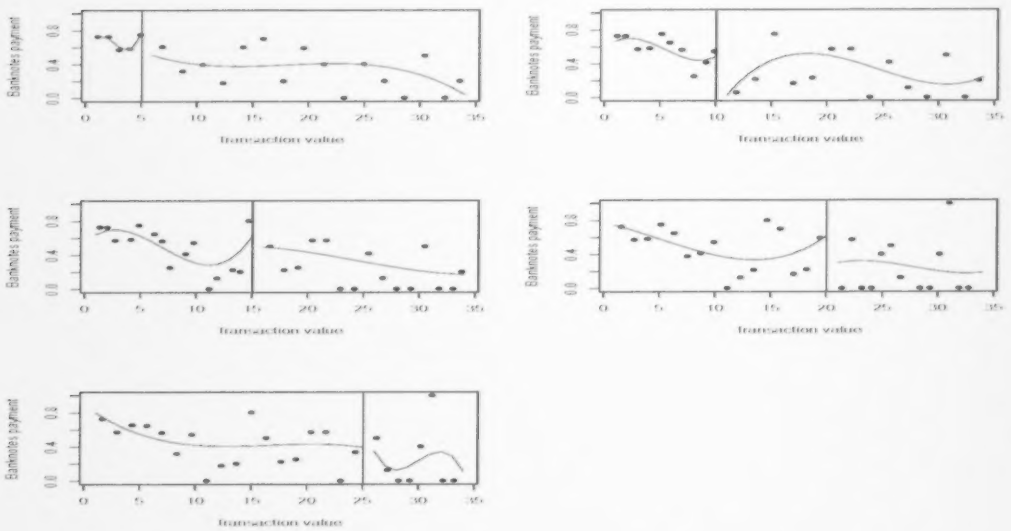
Note: Connecting lines do not reflect possible transactions.

Figure 5: Total (small and large) number of coins received as change for transactions between $0.00 and $25.00



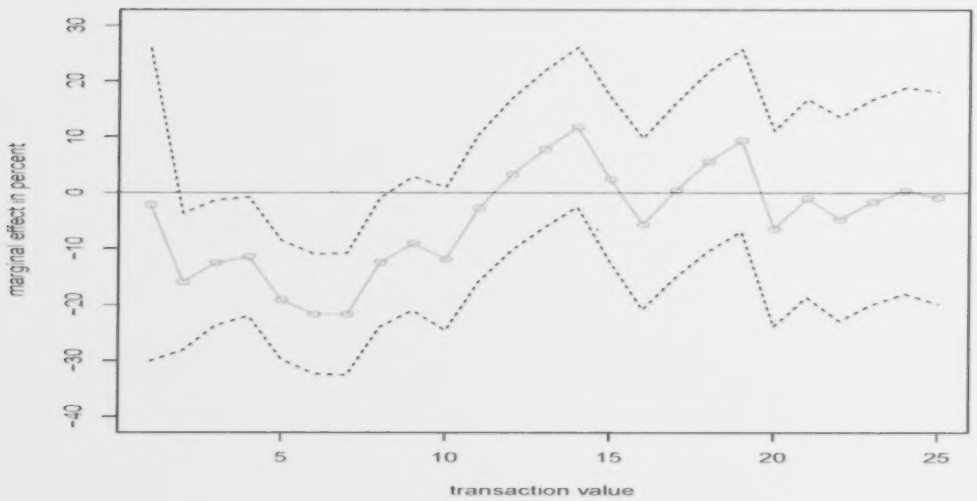
Notes: Red solid lines are approximated relationships between number of coins received as change and transaction values, where downward sloping curves in each five-dollar interval smooth out fast oscillations and small discontinuities.

Figure 6: Regression discontinuity plots for five thresholds



Notes: Regression discontinuity plots are generated by R command *rdplot* (Calonico, Cattaneo, and Titiunik (2015)). The evenly spaced binning is used, and thresholds are chosen to be 5.05, 10.05, 15.05, 20.05 and 25.05. Notice that such nonparametric plots are only illustrating the potential payment choice discontinuities, and should not be taken for providing the precise estimates.

Figure 7: Individual threshold regressions at each dollar interval



Notes: The red dots are coefficients of discontinuities in cash-use probabilities from 25 individual threshold regressions where each threshold is determined at every one-dollar transaction value. The dotted curves are the 95 percent confidence interval based on Lee and Card (2008).